Analyzing the convergence factor of residual inverse iteration

Elias Jarlebring
Wim Michiels

Report TW 570, July 2010

Katholieke Universiteit Leuven
Department of Computer Science
Celestijnenlaan 200A – B-3001 Heverlee (Belgium)

Analyzing the convergence factor of residual inverse iteration

Elias Jarlebring
Wim Michiels

Report TW 570, July 2010

Department of Computer Science, K.U.Leuven

Abstract

A complete characterization of the convergence factor can be very useful when analyzing the asymptotic convergence of an iterative method. We will here establish a formula for the convergence factor of the method called *residual inverse iteration*, which is a method for nonlinear eigenvalue problems and a generalization of the well known *inverse iteration*. The formula for the convergence factor is explicit and only involves quantities associated with the eigenvalue the iteration converges to, in particular the eigenvalue and eigenvector.

Besides deriving the explicit formula we also use the formula to characterize the convergence of the method. In particular, we derive a formula for the first order expansion when the shift is close to the eigenvalue. The residual inverse iteration allows some freedom in the choice of a vector r_k. In the analysis we characterize the convergence for different choices of r_k. We use the explicit formula for the first order expansion to show that the convergence factor approaches zero when the shift approaches the eigenvalue for an arbitrary choice of r_k. Moreover, we show that using an approximation of the left eigenvector as r_k as proposed in the literature, is natural since it results in accurate eigenvalue approximations; but it is not necessarily optimal in terms of the first order expansion of the convergence factor when the shift is close to the eigenvalue. The convergence factor also allows us to completely characterize the behavior of the method for double eigenvalues. For non-semisimple double eigenvalues it turns out that the convergence factor is one, implying slow or no convergence at all. For the case of a semisimple eigenvalue, the method behaves in a similar as for a simple eigenvalue, except that when converging, the iterates converge to the subspace of eigenvectors, not necessarily to a particular eigenvector.

manuscript No.
(will be inserted by the editor)

Analyzing the convergence factor of residual inverse iteration

Elias Jarlebring · **Wim Michiels**

Received: date / Accepted: date

Abstract A complete characterization of the convergence factor can be very useful when analyzing the asymptotic convergence of an iterative method. We will here establish a formula for the convergence factor of the method called *residual inverse iteration*, which is a method for nonlinear eigenvalue problems and a generalization of the well known *inverse iteration*. The formula for the convergence factor is explicit and only involves quantities associated with the eigenvalue the iteration converges to, in particular the eigenvalue and eigenvector.

Besides deriving the explicit formula we also use the formula to characterize the convergence of the method. In particular, we derive a formula for the first order expansion when the shift is close to the eigenvalue. The residual inverse iteration allows some freedom in the choice of a vector r_k. In the analysis we characterize the convergence for different choices of r_k. We use the explicit formula for the first order expansion to show that the convergence factor approaches zero when the shift approaches the eigenvalue for an arbitrary choice of r_k. Moreover, we show that using an approximation of the left eigenvector as r_k as proposed in the literature, is natural since it results in accurate eigenvalue approximations; but it is not necessarily optimal in terms of the first order expansion of the convergence factor when the shift is close to the eigenvalue. The convergence factor also allows us to completely characterize the behavior of the method for double eigenvalues. For non-semisimple double eigenvalues it turns out that the convergence factor is one, implying slow or no convergence at all. For the case of a semisimple eigenvalue, the method behaves in a similar as for a simple eigenvalue, except that when converging, the iterates converge to the subspace of eigenvectors, not necessarily to a particular eigenvector.

Keywords Nonlinear eigenvalue problems · Residual inverse iteration · Convergence factors

Mathematics Subject Classification (2000) 65F15 · 65H17 · 15A18

1 Introduction

Suppose $M : \mathbb{C} \to \mathbb{C}^{n\times n}$ corresponds to a matrix depending in a sufficiently smooth way on the parameter λ and consider the problem of finding $\lambda_* \in \mathbb{C}$ and $v_* \in \mathbb{C}\backslash\{0\}$ such that,

$$M(\lambda_*)v_* = 0. \tag{1.1}$$

This equation is sometimes called the nonlinear eigenvalue problem and includes many of the fundamental problems of numerical linear algebra as special cases. See, e.g. the survey papers [15,20] and the more recent problem collection [4].

Newton's method can be applied to the nonlinear equation (1.1). After some manipulations (often attributed to [23]) of the Newton iteration, it can be written as an iteration in n variables. The computationally dominating part in this Newton iteration is the solving of a linear system, involving a matrix depending on the eigenvalue iterate. The fact that the matrix depends on the iterate prevents efficient use of a factorization computed before the iteration starts and is the motivation for the method *residual inverse iteration* [16] where a linear system also has to be solved in each step, but the matrix can be kept constant, allowing us to use a factorization computed before the iteration starts.

Residual inverse iteration can be summarized as follows. See [16] for a thorough introduction. Given an approximation of the eigenvector v_k and a vector r_k (further discussed later) compute the solution to the scalar nonlinear equation

$$r_k^H M(\lambda)v_k = 0. \tag{1.2}$$

We assign the next eigenvalue iterate the value of a solution to (1.2), i.e., $\lambda_{k+1} \leftarrow \lambda$, and compute a new eigenvector approximation v_{k+1},

$$v_{k+1} = \alpha_k(v_k - M(\sigma)^{-1}M(\lambda_{k+1})v_k), \tag{1.3}$$

where the normalization constant α_k is chosen such that $c^H v_{k+1} = 1$. The complex number $\sigma \in \mathbb{C}$ is called the shift and is considered given by the user. The iteration consists of iterating the two steps corresponding to (1.2), i.e., solving the scalar equation, and (1.3), i.e., updating the eigenvector approximation.

Several choices for the vector r_k have been proposed. In [16], Neumaier proposes to set

i) $r_k^H := c^H M(\sigma)^{-1}$ if the problem has no particular symmetry; and
ii) $r_k := v_k$ if $M(\lambda_*)$ is Hermitian and $\lambda_* \in \mathbb{R}$.

If the eigenvalue problem is linear, these choices are natural since the method reduces to standard inverse iteration where r_k should be chosen as an approximation of the left eigenvector. Schreiber [21] also proposes to use a different type of approximation of the left eigenvector. Although the choice is natural for linear case, the understanding is not complete in the general situation. In order to constructively study how r_k should be chosen, we derive the theory for an arbitrary choice

$$r_k := r(v_k) \to r(v_*) =: r_*.$$

The iteration converges linearly to simple eigenvalues. Let $\Delta v_k := v_k - v_*$ and $\Delta\lambda_k := \lambda_k - \lambda_*$ denote the error in the eigenvector and eigenvalue correspondingly. We will see (in Section 2) that the iteration can be naturally treated as an iteration only in approximations of the eigenvectors. Hence, we will first consider the error in the eigenvector. We prove an asymptotic equations of the error of the eigenvector

$$\Delta v_{k+1} = B\Delta v_k + H.O.T.,$$

where B is a matrix depending on quantities involving the eigenvalue to which it converges and the notation *H.O.T.* as abbreviation for *higher order terms*. Since the generic situation is that the iteration speed is limited by the larges eigenvalue of B, we will call $\rho(B)$ the *convergence factor* for the iteration. Unlike other analysis of residual inverse iteration, including the original paper [16], we will here present a complete characterization of the convergence factor and use it to study the local convergence properties.

The property that the matrix $M(\sigma)$ in (1.3) can be kept constant throughout the iteration and the fact that the scalar equation (1.2) is computationally easy to solve in many situations, has made it attractive for many large-scale applications. It has been successfully combined with and used in modern methods for nonlinear eigenvalue problems [14,8]. The relation with the nonlinear Arnoldi method [24] is particularly noteworthy, since the residual inverse iteration is the motivation for the subspace expansion in [24] and it has been successfully used to solve many different types of nonlinear eigenvalue problems; see, e.g., [26,3,2,25] and related works. Residual inverse iteration has been used in [19] to study the stability of a time-delay system with periodic coefficients. An extended version of the method in [24] was presented in [13] and a two-sided variation is given and analyzed in [21, Section 4.2.2].

Despite the extensive use, no progress has to our knowledge been made in terms of qualitative understanding of the convergence, apart from the original paper [16] and some notes in [21]. Several aspects of the convergence remain open. Neumaier showed that a convergence basin exists for σ sufficiently close to the solution and that that it is linear convergence for simple eigenvalues with a convergence factor proportional to $\sigma - \lambda_*$ (for sufficiently small $\sigma - \lambda_*$) if r_k is chosen as an approximation of the eigenvector described above. Neumaier conjectures by experiments that the method also works for double eigenvalues. However, observations in an example Rott [19] indicates that the method can indeed fail for double eigenvalues. Schreiber [21] raises questions about the choice of r_k. Among other things, these open issues related to convergence to double eigenvalues and issues related to the choice of r_k will be addressed in this paper.

The main contributions of this paper can be summarized as follows. We derive an explicit expression for the convergence factor for simple eigenvalues (Section 3) and use it to analyze the situation where σ is close to the eigenvalue λ_*. In particular, we derive (in Section 4) a formula for the first order expansion when σ is close to λ_* and see that for any choice of r_k, the convergence factor approach zero as σ approach λ_*. If r_k approach the left eigenvector, then the convergence factor corresponding to the eigenvalue error is the square of the convergence factor corresponding to the eigenvector. In this sense, Neumaier's choice of r_k is good. We show by example that

the choice of r_k proposed by Neumaier is however necessarily the optimal choice in terms of growth rate of the convergence factor.

We also use the formula for the convergence factor to give a complete explanation of the convergence to double eigenvalues. In Section 6 we show that for non-semisimple eigenvalues we can expect that the method does not work or at least that the convergence is slow. For semisimple eigenvalues we illustrate that when converging, the convergence is linear and it converges to the span of the eigenvectors and not necessarily to a particular eigenvector.

Note that the convergence theory in the original paper [16] does not involve explicit expressions for the convergence factor. By using explicit expression for convergence factors we can in this paper go further and explain the convergence for double eigenvalues and characterize the growth rate of the convergence factor when λ_* is close to σ for an arbitrary choice of r_k.

There are a number of variations of Newton's method for nonlinear eigenvalue problems, e.g., the method of successive linear problems [20], block Newton [11], Rayleigh function iteration [22] and an approach by Kublanovskaya [12]. Some of the convergence properties of Newton's method (for nonlinear eigenvalue problems) are available [17, Section 5] and [1] including convergence factor analysis in [9]. Although these methods, as well as residual inverse iteration, are all somehow related to Newton's method, the derivation and analysis of the methods differ substantially and we have not found any direct use of the analysis of these methods in this work.

2 A fixed point formulation

In order to characterize the convergence in the following sections we will need to formulate the iteration in a compact way. The solution to the nonlinear scalar equation (1.2) will for notational purposes be denoted p. That is, $p : \mathbb{C}^n \to \mathbb{C}$ is the function fulfilling

$$r(v)^H M(p(v))v = 0,$$

and it is known as the generalized (nonlinear) Rayleigh functional. We will, as usual, make the genericy assumption that the Rayleigh functional locally defines a unique

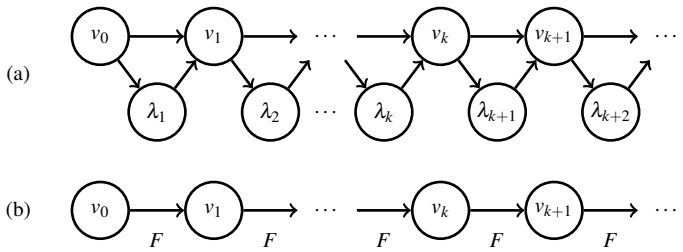


Fig. 2.1 The computation tree for residual inverse iteration (part (a)) and the fixed point reformulation (2.2) is shown in part (b).

solution. A sufficient condition that there locally exists a unique local solution is that

$$r_*^H M'(\lambda_*)v \neq 0. \tag{2.1}$$

See [18,6] and [21, Chapter 3] for more on nonlinear Rayleigh functionals.

The computation tree of the residual inverse iteration, i.e., (1.2) and (1.3), is visualized in Figure 2.1a. From the computation tree it is clear that, in order to compute λ_{k+1} only the value of v_k is needed. The only way λ_{k+1} depends on λ_k is via v_k. For this reason, it is for the purpose of asymptotic convergence analysis, more appropriate to reformulate the problem by eliminating λ_k. We can state the same iteration as an iteration of vectors of length n as can be seen in Figure 2.1b. More precisely, we represent the iteration by the fixed point iteration

$$v_{k+1} = F(v_k), \tag{2.2}$$

where

$$F(v) := \alpha(v)(v - M(\sigma)^{-1}M(p(v))v),$$

and $\alpha(v)$ is such that $c^H F(v) = 1$, i.e., $\alpha(v) := 1/c^H(v - M(\sigma)^{-1}M(p(v))v)$.

In the following result we show that the fixed points of the iteration (2.2) are indeed solutions to the nonlinear eigenvalue problem. Moreover, the fixed points are isolated if $M(p(v))$ has a null-space of dimension one, which is the case for simple eigenvalues. The case where the null-space has rank greater than one corresponds to a non-isolated fixed point and will be further analyzed in the section on double eigenvalues (Section 6).

Proposition 2.1 (Existence of a fixed point) *Consider a vector $v \in \mathbb{C}^n$ with $c^H v \neq 0$. Then the following statements are equivalent:*

i) $F(v) = v$.
ii) $M(p(v))v = 0$ *and* $c^H v = 1$.

Proof Suppose i) holds. Then, $F(v) = v$ and $1 = c^H F(v) = c^H v$. It follows that $\alpha(v) = 1$ and

$$v = F(v) = v - M(\sigma)^{-1}M(p(v))v,$$

from which we see that ii) holds. The converse follows analogously from direct manipulations.

3 The convergence factor

We are now ready to study the error of the residual inverse iteration using the fixed point formulation in the previous section. In order to derive a convergence factor we will need the following asymptotic relation between the eigenvalue error $\Delta\lambda_{k+1}$ and the eigenvector error Δv_k.

Lemma 3.1 *For sufficiently small eigenvector error Δv_k and eigenvalue error $\Delta\lambda_k$, the errors in the eigenvector are related by,*

$$\Delta\lambda_{k+1} = -\frac{r_*^H M(\lambda_*)\Delta v_k}{r_*^H M'(\lambda_*)v_*} + H.O.T. \tag{3.1}$$

Proof By forming the Taylor expansion for the condition that λ_{k+1} fulfills the Rayleigh function, i.e., $r_k^H M(\lambda_{k+1})v_k = 0$, we find that

$$0 = \left(r(v_*)^H + \Delta v_k^H J(v_*)^H\right)(M(\lambda_*) + \Delta\lambda_{k+1}M'(\lambda_*))(v_* + \Delta v_k) + H.O.T., \quad (3.2)$$

where $J(v_*) \in \mathbb{C}^{n\times n}$ is the Jacobian of $r : \mathbb{C}^n \to \mathbb{C}^n$.

The relation (3.1) follows by solving (3.2) for $\Delta\lambda_{k+1}$ and using that $M(\lambda_*)v_* = 0$. The proof is complete.

The relation now allows us to show a first order result for the eigenvector error. This will turn out to contain sufficient information for the analysis of the convergence factor.

Theorem 3.2 (The convergence factor) *For sufficiently small eigenvector error Δv_k, it fulfills the iteration,*

$$\Delta v_{k+1} = (I - v_* c^H)A\Delta v_k + H.O.T., \quad (3.3)$$

where

$$A := I - M(\sigma)^{-1}M(\lambda_*) + \frac{M(\sigma)^{-1}M'(\lambda_*)v_* r_*^H M(\lambda_*)}{r_*^H M'(\lambda_*)v_*}. \quad (3.4)$$

Proof We denote

$$z_k := v_k - M(\sigma)^{-1}M(\lambda_*)v_k$$

and neglect the higer order terms,

$$\begin{aligned} z_k &= v_* + \Delta v_k - M(\sigma)^{-1}M(\lambda_*)\Delta v_k - \Delta\lambda_{k+1}M(\sigma)^{-1}M'(\lambda_*)(v_* + \Delta v_k) + H.O.T. \\ &= v_* + A\Delta v_k + H.O.T., \end{aligned}$$

where we used (3.1) in the last equality.

The vector part of the iteration, i.e., the fixed point iteration F in (2.2), can now be written as $(c^H z_k)v_{k+1} = z_k$, i.e.,

$$c^H(v_* + A\Delta v_k)(v_* + \Delta v_{k+1}) = A\Delta v_k + H.O.T.$$

and since $c^H v_* = 1$,

$$\Delta v_{k+1} = (I - v_* c^H)A\Delta v_k + H.O.T.$$

We have completed the proof by showing (3.3).

It is now important to note that the first order characterization of the eigenvector error in (3.3) contains the information necessary to describe the asymptotic behavior of the error. The linear iteration (3.3) is essentially a power iteration and will therefore generically have a convergence factor equal to the modulus of the largest eigenvalue of the iteration matrix $(I - v_* c^H)A$. Hence, the generic situation is that the convergence factor of residual inverse iteration is

$$\frac{\|v_{k+1} - v_*\|_2}{\|v_k - v_*\|_2} \to \rho((I - v_* c^H)A), \quad (3.5)$$

where $\rho : \mathbb{C}^{n\times n} \to \mathbb{R}$ denotes the spectral radius, i.e., the modulus of the largest eigenvalue. The above argument can be formalized by noting that (3.3) is the linearization of the fixed point iteration $v_{k+1} = F(v_k)$ around the fixed point $v_* = F(v_*)$. With the formalized argument involving Banach's contraction mapping theorem one finds a sufficient (and possibly necessary) conditions for the existance of an attractive fixed point. In this paper on convergence factors, it will be sufficient to note that the linearized equation (3.3) has the convergence factor given by (3.5) which is generically the convergence factor for the whole iteration.

Remark 3.1 (The matrix A and the iteration matrix) The construction of the iteration matrix

$$(I - v_* c^H)A$$

can be interpreted as follows. The operation of multiplying the matrix A from the left by $(I - v_* c^H)$ only changes one of the eigenvalues of A. Note that the matrix A always has an eigenvalue equal to one since $Av_* = v_*$. Moreover, we have that $c^H v_* = 1$ and the multiplication from the left with $(I - v_* c^H)$ corresponds to transforming the eigenvalue of v_* to zero and leaving all other eigenvalues unchanged. Hence, the eigenvalues of $(I - v_* c^H)A$ are the same as the eigenvalues of A except for the eigenvalue 1 corresponding to eigenvector v.

This is consistent with the fact that the iteration is independent of c (as long as $c^H v_k \neq 0$). Although the iteration matrix in (3.3) depends on c, the convergence factor given by (3.5) is also an eigenvalue of A which is independent of c.

The convergence factor of the fixed point iteration (2.2) is completely characterized by (3.5). Since the fixed point iteration is an iteration in the vectors v_k, the convergence factor (3.5) does not directly imply anything about the convergence factor corresponding eigenvalue iterates. We will now characterize eigenvalue iterates also using the matrix A.

For the eigenvalue convergence factor we need to separate between the case if $r_*^H M(\lambda_*) = 0$ or not, i.e., if r_* is a left eigenvector. We characterize the case $r_*^H M(\lambda_*) \neq 0$ in Corollary 3.3 and the situation $r_*^H M(\lambda_*) = 0$ in Remark 3.2.

Suppose $r_*^H M(\lambda_*) \neq 0$. Then the eigenvalue convergence factor is equal to the convergence factor of the fixed point iteration, in the following sense.

Corollary 3.3 (Eigenvalue convergence factor) *Suppose $r_*^H M(\lambda_*) \neq 0$. The eigenvalue convergence factor is determined by*

$$\frac{\Delta\lambda_{k+1}}{\Delta\lambda_k} = \frac{r_*^H M(\lambda_*) A \Delta v_k}{r_*^H M(\lambda_*) \Delta v_k} + H.O.T.$$

Moreover, if the iteration converges to a simple eigenvalue, then the generic situation is that

$$\frac{|\Delta\lambda_{k+1}|}{|\Delta\lambda_k|} \to \rho((I - v_* c^H)A).$$

Proof From Lemma 3.1 it follows that

$$\frac{\Delta\lambda_{k+1}}{\Delta\lambda_k} = \frac{r_*^H M(\lambda_*) \Delta v_{k+1}}{r_*^H M(\lambda_*) \Delta v_k} + H.O.T.$$

Theorem 3.2 implies that $\Delta v_{k+1} = (I - v_* c^H)A\Delta v_k + H.O.T.$, and Δv_k approach a right eigenvector of $(I - v_* c^H)A$. The proof is completed noting that $r_*^H M(\lambda_*)\Delta v_k \to 0$ is a degenarate situation.

Remark 3.2 (Eigenvalue convergence factor if $r_*^H M(\lambda_*) = 0$) If r_* is a left eigenvector we have that $r_*^H M(\lambda_*)\Delta v_k = 0$. This is the case in particular when $M(\lambda_*)$ is Hermitian and λ_* real and we set (as Neumaier proposes) $r_k = v_k$. For this case we note that the relation between the eigenvalue error and the eigenvector error in Lemma 3.1 can be changed to

$$\Delta\lambda_{k+1} = -\frac{\Delta v_k^H J(v_*)^H M(\lambda_*)\Delta v_k}{r_*^H M'(\lambda_*)v_*} + H.O.T., \tag{3.6}$$

where $J(v_*)$ is the Jacobian of $r : \mathbb{C}^n \to \mathbb{C}^n$. The relation (3.6) has the interpretation that the eigenvalue is twice as accurate as the eigenvector, in the sense that $|\Delta\lambda_{k+1}| \sim \|\Delta v_k\|_2^2$. More precisely,

$$\frac{|\Delta\lambda_{k+1}|}{|\Delta\lambda_k|} = \frac{|\Delta v_k^H J(v_*)^H M(\lambda_*)\Delta v_k|}{|\Delta v_{k-1}^H J(v_*)^H M(\lambda_*)\Delta v_{k-1}|} + H.O.T. =$$
$$\frac{|\Delta v_{k-1}^H A^H (I - v_* c^H)^H J(v_*)^H M(\lambda_*)(I - v_* c^H)A\Delta v_{k-1}|}{|\Delta v_{k-1}^H J(v_*)^H M(\lambda_*)\Delta v_{k-1}|} + H.O.T. \to$$
$$\left(\rho((I - v_* c^H)A)\right)^2,$$

where in the last step we used that Δv_k approach the eigenvector of $(I - v_* c^H)A$ corresponding to the largest eigenvalue and assumed that $\Delta v_k^H J(v_*)^H M(\lambda_*)\Delta v_k \not\to 0$. Note that twice as accurate in this sense does not imply quadratic convergence, but squaring the convergence factor.

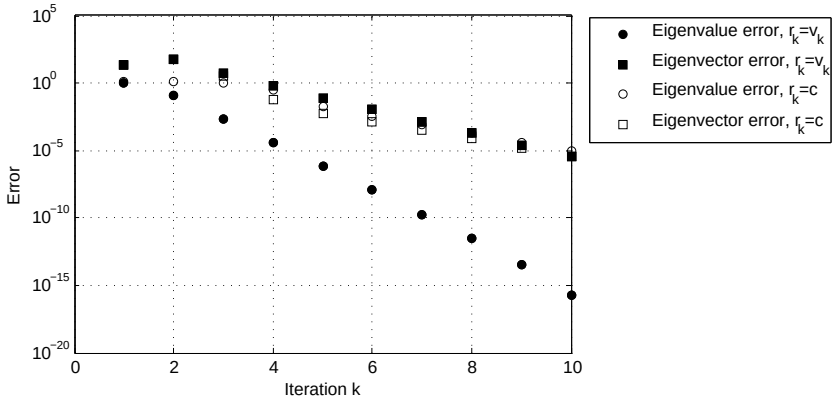


Fig. 3.1 Convergence for real eigenvalue of the nonlinear eigenvalue problem (3.7).

As an illustration we consider the nonlinear eigenvalue problem

$$M(\lambda) = -\lambda I + \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 3 \end{pmatrix} + \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & 6 \end{pmatrix} e^{-\lambda}, \tag{3.7}$$

which is symmetric and has a real eigenvalue $\lambda_* \approx 3.18581$. We carry out residual inverse iteration with $c = (1,1,1)^T$ and $\sigma = 3$, for different choices of r_k. The error is illustrated in Figure 3.1. We clearly see that the convergence factor for the eigenvalue is the essentially the squared of the convergence factor for the eigenvector if $r_k = v_k$. The eigenvector convergence does however not appear faster than for the choice $r_k = c$. Note that the convergence is linear for any choice of r_k, which clarifies the convergence order concers mentioned in [21, page 96].

4 Convergence factor for shift close to the eigenvalue

Due to the interpretation of σ as a shift, it is to expect that if it is close to an eigenvalue, we have fast convergence. This was already proven in Neumaier [16, page 919] for the choices of the vector $r_k^H = c^H M(\sigma)^{-1}$ and $r_k = v_k$, where it is shown that the convergence factor grows linearly with the distance between the shift and the eigenvalue. We will in this section see that this statement holds for arbitrary r_k.

Moreover, we derive explicit formulas for the growth of convergence factor as a function of the shift-eigenvalue distance. With the explicit formulas we can also characterize what is an optimal choice of r_k in sense of growth rate, which by example is shown to not necessarily be the left eigenvector.

In order to derive the main result we need the following lemma. The proof is available in Appendix A.

Lemma 4.1 *Suppose σ is close to λ_* and suppose $M(\lambda_*)$ has a null space of dimension one. Let u and v left and right null-vectors of $M(\lambda_*)$ and suppose $u^H M'(\lambda_*)v \neq 0$. Then,*

$$M(\lambda_*)M(\sigma)^{-1} = A_0 + (\sigma - \lambda_*)A_1 + O(\sigma - \lambda_*)^2, \tag{4.1}$$

where $A_0 \in \mathbb{C}^{n\times n}$ is given by

$$A_0 = I - \frac{M'(\lambda_*)vu^H}{u^H M'(\lambda_*)v}$$

and $A_1 \in \mathbb{C}^{n\times n}$ is

$$A_1 = A_0\left(-M'(\lambda_*)M(\lambda_*)^+ A_0 - \frac{M''(\lambda_*)vu^H}{u^H M'(\lambda_*)v}\right)\left(I + \frac{M'(\lambda_*)vu^H}{u^H M'(\lambda_*)v}\right)^{-1},$$

where $(\cdot)^+$ denotes the pseudo inverse.

It turns out that when the shift is close to the eigenvalue, the convergence factor is also small.

Theorem 4.2 (Shift close to the eigenvalue) *When σ is close to the simple eigenvalue λ_*, then the convergence factor is to first order given by,*

$$\rho((I - v_* c^H)A) = |\sigma - \lambda_*|\rho\left(\left(I - \frac{M'(\lambda_*)v_* r_*^H}{r_*^H M'(\lambda_*)v_*}\right)A_1\right) + O(\sigma - \lambda_*)^2 \tag{4.2}$$

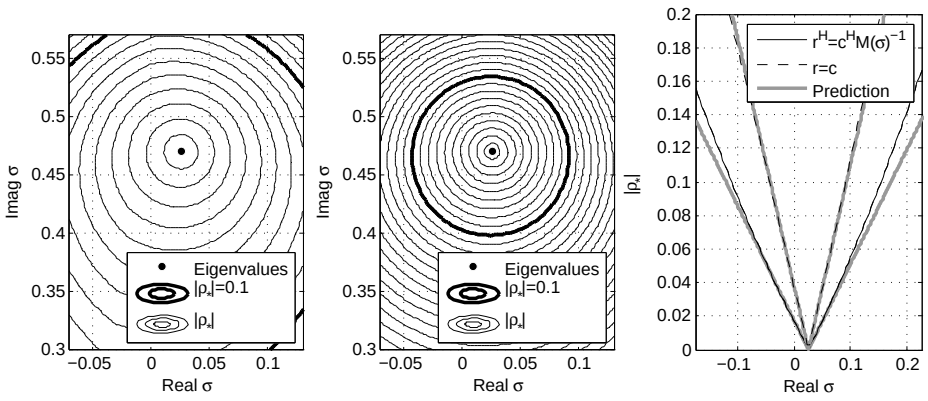
and in particular if $r_* = u_*$,

$$\rho((I - v_* c^H)A) = |\sigma - \lambda_*|\rho(A_1) + O(\sigma - \lambda_*)^2 \tag{4.3}$$

Proof Suppose $M(\lambda_*)$ has a one dimensional left and a one dimensional right nullspace, which are represented by the (normalized) eigenvectors u and v.

We will need the following relation several times,

$$\begin{aligned}(I - \frac{M'(\lambda_*)vr^H}{r^H M'(\lambda_*)v})A_0 = \\ I - M'(\lambda_*)v\left(\frac{r^H}{r^H M'(\lambda_*)v} + \frac{u^H}{u^H M'(\lambda_*)v} - \frac{r^H}{r^H M'(\lambda_*)v}\frac{M'(\lambda_*)vu^H}{u^H M'(\lambda_*)v}\right) = \\ I - \frac{M'(\lambda_*)vr^H}{r^H M'(\lambda_*)v}.\end{aligned} \tag{4.4}$$



(a) $r_k^H = c^H M(\sigma)^{-1}$

(b) $r_k = c = (1,1,1)^T$

(c) Predicted and observed growth rate of the convergence factor.

Fig. 4.1 Illustration of the growth rate of the convergence factor when the shift σ is close to the eigenvalue λ_* for the nonlinear eigenvalue problem (4.6).

Consider the following similarity transformation of A,

$$M(\sigma)AM(\sigma)^{-1} = I - \left(I - \frac{M'(\lambda_*)v_* r_*^H}{r_*^H M'(\lambda_*)v_*}\right) M(\lambda_*)M(\sigma)^{-1} =$$
$$I - \left(I - \frac{M'(\lambda_*)v_* r_*^H}{r_*^H M'(\lambda_*)v_*}\right) \left(I + (\sigma-\lambda_*)\left(-M'(\lambda_*)M(\lambda_*)^+ A_0\right.\right.$$
$$\left.\left. - \frac{M''(\lambda_*)v_* u^H}{u^H M'(\lambda_*)v_*}\right)\left(I + \frac{M'(\lambda_*)v_* u^H}{u^H M'(\lambda_*)v_*}\right)^{-1}\right) =$$
$$\frac{M'(\lambda_*)v_* r_*^H}{r_*^H M'(\lambda_*)v_*} -$$
$$(\sigma-\lambda_*)\left(I - \frac{M'(\lambda_*)v_* r_*^H}{r_*^H M'(\lambda_*)v_*}\right)\left(-M'(\lambda_*)M(\lambda_*)^+ A_0 - \frac{M''(\lambda_*)v_* u^H}{u^H M'(\lambda_*)v_*}\right)\left(I + \frac{M'(\lambda_*)v_* u^H}{u^H M'(\lambda_*)v_*}\right)^{-1} \tag{4.5}$$

Now note that according to Remark 3.1 we should remove the eigenvalue of A corresponding to the eigenvalue one. This corresponds (in (4.5)) to removing the term $M'(\lambda_*)vr^H/r^H M'(\lambda_*)v$. We have shown (4.2). The formula (4.3) is proven analously by directly applying Lemma 4.1.

Remark 4.1 (The choice of r_k) Neumaier [16] and Schreiber [21] propose to choose r_k as an approximation of the left eigenvector. Although this is a natural choice for linear case (inverse iteration) and the eigenvalue convergence factor is expected to be small (see Remark 3.2), we will now see that it is not necessarily optimal in terms of growth rate. Consider the following cubic polynomial eigenvalue problem,

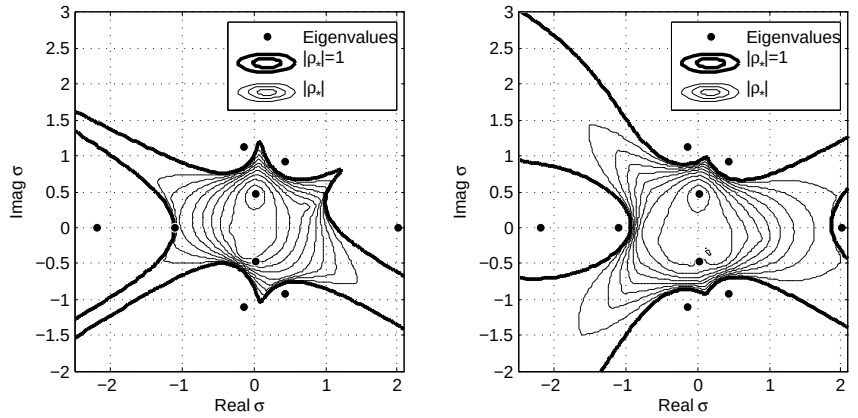
$$M(\lambda) := \begin{pmatrix} -16 & -4 & 7 \\ -14 & 7 & 13 \\ 6 & 8 & 7 \end{pmatrix} + \lambda^2 \begin{pmatrix} 2 & -6 & 1 \\ -2 & 22 & 11 \\ 7 & -1 & 1 \end{pmatrix} + \lambda^3 \begin{pmatrix} -4 & 3 & 12 \\ -17 & -11 & 0 \\ 1 & -1 & 3 \end{pmatrix}. \tag{4.6}$$

In Figure 4.1, where we illustrate the convergence factor and the growth rate for this example, we see the growth rate for $c = (1,1,1)$ and for $r_k^H = c^H M(\sigma)^{-1}$.

The growth rate for $r_k = u$ is $\rho' \approx 0.685$. With a simple optimization procedure we find a better growth rate $\rho' \approx 0.658$ for $r_k = (0.630, -0.754, 0.185)^T$. This contradicts the idea that Neumaiers choice is optimal in terms of growth rate of convergence factor. We note that at least for this example, the difference between the found optimum and the growth rate corresponding to the left eigenvector is small.

5 Convergence basin as a function of the shift

We saw above that the formula for the convergence factor in Section 3 could be used to characterize the convergence factor when the shift-eigenvalue distance is small. We will now illustrate the use of the formula for the convergence factor not only locally. In this non-local analysis we can also illustrate the different choices of r_k.



(a) $r^H = c^H M(\sigma)^{-1}$ (b) $r = u$

Fig. 5.1 Level contours of the convergence factor as a function of shift σ. The level sepation is 0.1.

We again consider the nonlinear eigenvalue problem (4.6). The convergence factor corresponding to the case where the iteration is started sufficiently close to $\lambda_* \approx 0.5i$ is shown in Figure 5.1 for two different choices of r_k. The left subfigure (Figure 5.1a) shows the convergence factor for the choice proposed by Neumaier. The convergence factor when r_k is chosen as the left eigenvector is illustrated in the right subfigure (Figure 5.1b). The region enclosed by the bold curve, i.e., a convergence factor equal to one, corresponds to possible shifts where the iteration will converge if the starting value is sufficiently close to $\lambda_* \approx 0.5i$.

The figure indicates that in this case, the acceptable choices of σ is larger if r_k is close to a left eigenvector, supporting Schreiber's suggestion [21, page 82] that an approximation of the left eigenvector is a good choice. Note however that this example is not strictly conclusive, since there is a region of the complex plane where the choice of Neumaier (Figure 5.1a) is acceptable but the convergence factor corresponding to the left eigenvector is not.

6 Double eigenvalues

In order characterize the situation where λ_* is a double eigenvalue we will use the concepts and generalizations of Jordan chains and generalized eigenvectors for nonlinear eigenvalue problems used in [5, Section 1.4] and [7]. The classification can be summarized as follows. A double eigenvalue can be either semisimple or non-semisimple. The matrix $M(\lambda_*)$ has a null-space of dimension (exactly) two if and only if the eigenvalue is semisimple. Correspondingly, $M(\lambda_*)$ has a null-space of dimension (exactly) one if the double eigenvalue is non-semisimple. Moreover, a double eigenvalue is non-semisimple if and only if there is a vector $v \in \mathbb{C}^n$ called a

generalized eigenvector such that

$$M'(\lambda_*)v_* = M(\lambda_*)\dot{v}. \tag{6.1}$$

6.1 A double non-semisimple eigenvalue

Suppose λ_* is a double non-semisimple eigenvalue, which is the generic situation for double eigenvalues of nonlinear eigenvalue problem without any particular structure. Note that this does not change the fixed point formulation (in Section 2). In particular, we still have from Proposition 2.1 that the fixed point corresponding to a double eigenvalue exists and is isolated.

Suppose r_* is not a left eigenvector. We know from Remark 3.1 that the matrix A always has an eigenvalue one corresponding to eigenvector v. It is now easy to verify that $\dot{v}$ is also an eigenvector of A corresponding to the eigenvalue one, i.e., one is a double eigenvalue of A. It follows from the definition of the generalized eigenvector (6.1) that,

$$A\dot{v} = \dot{v} - M(\sigma)^{-1}M(\lambda_*)\dot{v} + \frac{M(\sigma)^{-1}M'(\lambda_*)v_* r_*^H M(\lambda_*)\dot{v}}{r_*^H M'(\lambda_*)v_*} = \dot{v}. \tag{6.2}$$

The iteration matrix $(I - v_* c^H)A$ has the same eigenvalues as A except for the eigenvalue corresponding to v_* which is transformed to zero (see Remark 3.1). Loosely speaking, we now have that one of the double eigenvalues of A is removed when instead considering $(I - v_* c^H)A$, but one remains. The eigenvector corresponding to eigenvalue one is explicitly $\hat{v} := \dot{v} - (c^H\dot{v})v_*$. That is,

$$(I - v_* c^H)A\hat{v} = (I - v_* c^H)A(\dot{v} - (c^H\dot{v})v_*) = (I - v_* c^H)A\dot{v} = (I - v_* c^H)\dot{v} = \hat{v}. \tag{6.3}$$

From the fact that the iteration matrix $(I - v_* c^H)A$ always has an eigenvalue one we predict that for double non-semisimple eigenvalues the convergence will be slow or the iteration will not converge at all. This is consistent with the application of residual inverse iteration in [19], where one bifurcation curve, which corresponds to a double eigenvalue, can not be accurately followed.

We here assumed that r_* is not a left eigenvector. If this is the case, the iteration is no longer well posed since the Rayleigh functional $p(\cdot)$ no longer uniformly defines a unique solution. That is, the condition (2.1) is violated. Once v_k is close to the solution, the Rayleigh functional will have (at least) two solutions close to the exact eigenvalues λ_* and the iteration is does not uniquely define a next eigenvalue iterate.

Example 6.1 (Double non-semisimple eigenvalue) Consider as in [10],

$$M(\lambda) = -\lambda I + A_0 + A_1 e^{-\lambda},$$

where

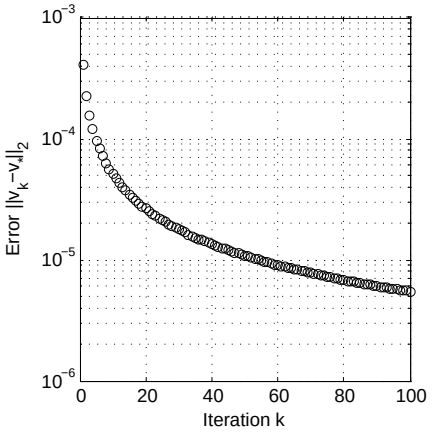
$$A_0 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_3 & -a_2 & -a_1 \end{pmatrix}, A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -b_3 & -b_2 & -b_1 \end{pmatrix},$$

and

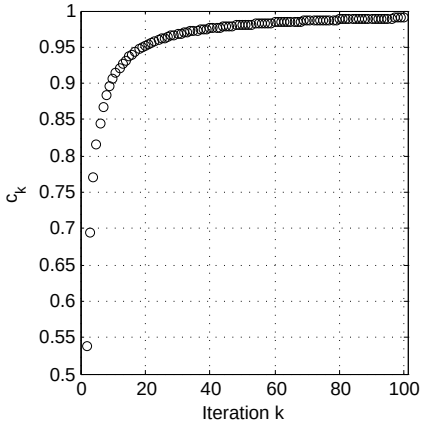
$$a_1 = \frac{2}{5}\frac{(65\pi+32)}{8+5\pi} \approx 3.98,\ a_2 = \frac{9\pi^2(13+5\pi)}{8+5\pi} \approx 108,$$
$$a_3 = \frac{324}{5}\frac{\pi^2(5\pi+4)}{8+5\pi} \approx 531,\ b_1 = \frac{260\pi+128+225\pi^2}{10(8+5\pi)} \approx 13.6,$$
$$b_2 = \frac{45\pi^2}{10(8+5\pi)} \approx 18.7 \text{ and } b_3 = \frac{81\pi^2(40\pi+32+25\pi^2)}{10(8+5\pi)} \approx 1363.$$

This nonlinear eigenvalue problem has a double non-semisimple eigenvalue for $\lambda = 3\pi i$.

We apply residual inverse iteration to this problem with the paramters $c = (1,1,1)^T$, $\sigma = 9.4i$, $r = M(\sigma)^{-H}c$ and $v_0 = v_* + (10^{-3}, 10^{-3}, 10^{-3})^T$. In the convergence diagram in Figure 6.1a we see that the convergence stagnates although the shift as well as the starting value is very close to the exact solution. Figure 6.1b shows the estimate of the convergence factor $c_k := \|v_{k+1} - v_*\|/\|v_{k+1} - v_*\|$. In this figure (Figure 6.1b), it is clear that it approaches one, as predicted by (6.3).



(a) The convergence is slow for the double eigenvalue.



(b) The factor $c_k := \|v_{k+1} - v_*\|_2/\|v_k - v_*\|_2$ approaches 1.

Fig. 6.1 Illustration of the convergence for the example with a double non-semisimple eigenvalue (Example 6.1).

6.2 A double semisimple eigenvalue

Now suppose λ_* is double semisimple eigenvalue, i.e., the null space of $M(\lambda_*)$ has dimension two. Let v_1 and v_2 be a basis such that

$$M(\lambda_*)v_1,\ M(\lambda_*)v_2 = 0,\ v_1 \neq v_2,\ v_1, v_2 \in \mathbb{C}^n \setminus \{0\}.$$

Since any linear combination of v_1 and v_2 is also an eigenvector, it now follows from Proposition 2.1 that the fixed point of F is (unlike the non-semisimple case) *not* isolated. More precisely, the fixed points of F form the set

$$\Omega := \left\{ v \in \mathbb{C}^n \setminus \{0\} :\ v \in \operatorname{span}\{v_1, v_2\},\ \ c^H v = 1 \right\}.$$

In the generic case, where either $c^H v_1 \neq 0$ or $c^H v_2 \neq 0$, Ω is a one-dimensional space.

Let $v_* \in \Omega$ and consider the corresponding spectrum of the iteration matrix

$$(I - v_* c^H)A. \tag{6.4}$$

Take $\tilde{v} \in \operatorname{span}\{v_1, v_2\}$ such that

$$c^H \tilde{v} = 0.$$

Note that such a vector can always be found under the non-degeneracy condition mentioned above. We have

$$(I - v_* c^H)A\tilde{v} = (I - v_* c^H)\tilde{v} = \tilde{v}.$$

Thus the iteration matrix (6.4) has an eigenvalue equal to one, with corresponding eigenvector $\tilde{v}$.

We have hence shown that similar to the non-semisimple case, A has a double eigenvalue at one. Moreover, also similar to the non-semisimple case, we have that the iteration matrix $(I - v_* c^H)A$ only has a simple eigenvalue at one. However, in contrast to non-semisimple case, the eigenvalue one of the iteration matrix for the semisimple case does *not* affect the convergence rate of the eigenvalues. This can be seen from the following argument. First note that the corresponding eigenvector $\tilde{v}$ is such that

$$(v_* + \varepsilon\tilde{v}) \in \Omega,\ \ \forall \varepsilon \in \mathbb{R}.$$

In words, a perturbation of the fixed point v_* in the direction of $\tilde{v}$, for which the linearized analysis is inconclusive about the recovery (eigenvalue one), corresponds to a perturbation *along* the line of fixed points. The conclusion of this reasoning is that for double semisimple eigenvalues one should study

$$\sigma((I - v_* c^H)A) \setminus \{1\} \tag{6.5}$$

in order to establish the convergence factor. Hence, the convergence factor depends on the fixed point, or equivalently, the eigenvector, under consideration. Because there is a one-dimensional subspace of (normalized) eigenvectors Ω, the method will, when converging, not converge to a *specific* eigenvector. In theory, the asymptotics of the iteration could also exhibit a drift along the subspace of eigenvectors Ω. The latter has however not been observed in our numerical experiments.

Example 6.2 (A double semisimple eigenvalue) We will now consider the following example in order to illustrate the linear convergence for a double semisimple eigenvalue. The nonlinear eigenvalue problem corresponding to

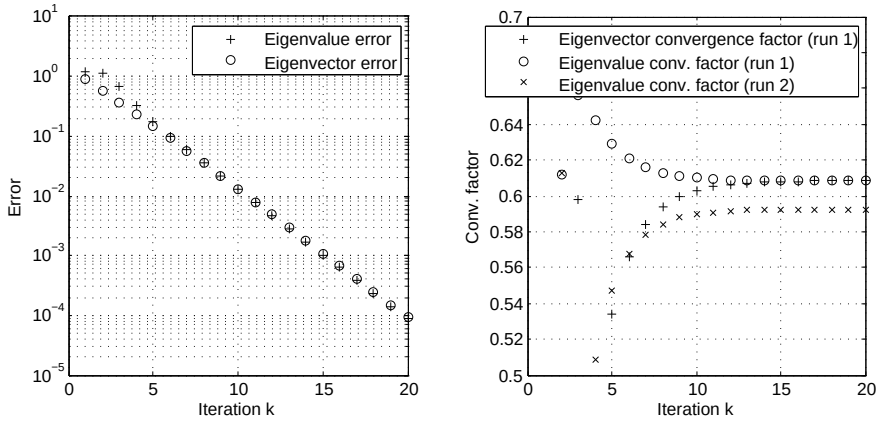
$$M(\lambda) := \begin{pmatrix} 1-\lambda^2 & 0 & \lambda \\ 0 & 2(1-\lambda^2) & 0 \\ 1-\lambda & 0 & \lambda^3+2\lambda^2+1 \end{pmatrix},$$

has a double eigenvalue for $\lambda = 1$. It is easy to verify that two corresponding right eigenvectors are $v = e_1$ and $v = e_2$. We fix $r = (1,1,2)^T$, $c = (1,1,1)^T$ and start the iteration with $v_0 = (1,2,1)^T$.

We first observe in Figure 6.2a that the convergence is linear. The figure shows the eigenvalue error $|\lambda_k - \lambda_*| = |\lambda_k - 1|$ and the difference $|v_k - v_K|$, where v_k is the eigenvector iterate and v_K an eigenvector iterate approximating a converged eigenvector ($K = 50$). We use this type of construction since, although the eigenvectors are $v = e_1$ and $v = e_2$, the iteration converges to a linear combination, which is not known a priori.

The estimate of the convergence factor is visualized in Figure 6.2b, for both the eigenvector approximation and the eigenvalue approximation, i.e., $(\lambda_{k+1} - \lambda_*)/(\lambda_k - \lambda_*)$ and $\|v_{k+1} - v_K\|_2 / \|v_k - v_K\|_2$.

In a second run, we start the iteration with a different vector $v_0 = \frac{1}{2}(2,1,2)^T$. Note in Figure 6.2b that convergence factor changes with starting value. Although the asymptotic eigenvector error in Theorem 3.2 holds as long as the eigenvector converges (which appears to be the case for this example), it depends on the eigenvector, which is not unique. Different eigenvectors of the manifold Ω will yield different A and hence different convergence factors.



(a) Error vs. iteration (run 1)

(b) Estimated convergence factor

Fig. 6.2 Visualization of the (linear) convergence for the semisimple eigenvalue in Example 6.2.

In general, the fixed point iteration does not necessarily converge to a specific eigenvector in the manifold. We observe convergence for this and other examples with semisimple eigenvalues and have illustrated that the iteration converges (in a sense) to the manifold Ω, and not necessarily to an element of Ω.

7 Conclusions and outlook

One of the main goals of this paper is to characterize the convergence of the method residual inverse iteration. The approach we take is based on an analysis using the convergence factor. We have derived an explicit formula for the convergence factor as well as an explicit formula for the first order expansion of the convergence factor when the shift is close to the eigenvalue. These formulas can be used to characterize properties of the method. In particular, we show that if r_k approach the left eigenvector, we can expect that the convergence factor of the eigenvalue is the square of the convergence factor of the eigenvector, indicating that Neumaier's choice of r_k is good. We also show that although the choice of r_k Neumaier proposes is natural, it is not always the best choice in terms of the convergence factor and in particular in term of the first order expansion of the convergence factor. With the first order expansion we also show that, independent of r_k, the convergence factor approach zero when the shift approach the eigenvalue.

The explicit formula for the convergence factor is also used to completely characterize the behavior of the method for double eigenvalues, including non-semisimple as well as semisimple. For non-semisimple double eigenvalues, the method is not expected to converge or at least have slow convergence, whereas for the semisimple case, when convergent, we have linear convergence and a behavior similar for a simple eigenvalue.

Finally, we also want to point out that the explicit formula for the convergence factor may be valuable by itself. If the convergence factor can be accurately and cheaply estimated, it can be used to accelerate the method or even increase the convergence order by predicting the error of the iterate, similar to the technique of extrapolation.

A The proof of Lemma 4.1

Throughout this derivation we will need the following two formulas. Suppose E is a singular matrix with a null-space of dimension one, with left and right null vectors u and v and $u^H v \neq 0$. Then,

$$\operatorname{adj}(E) = -\frac{\|\operatorname{adj}(E)\|_2}{u^H v} v u^H =: \beta v u^H. \tag{A.1}$$

We will also use the Jacobi formula for $\frac{d}{d\sigma}\det(M(\sigma))$,

$$\left(\frac{d}{d\sigma}\det(M(\sigma))\right)_{\sigma=\lambda_*} = \operatorname{Tr}(\operatorname{adj}(M(\lambda_*))M'(\lambda_*)) = \beta \operatorname{Tr}(v u^H M'(\lambda_*)) = \beta u^H M'(\lambda_*) v. \tag{A.2}$$

The quantity which expansion we seek is now,

$$M(\lambda_*)M(\sigma)^{-1} = I + (M(\lambda_*) - M(\sigma))\frac{\operatorname{adj}(M(\sigma))}{\det(M(\sigma))} = I - (\sigma-\lambda_*)M'(\lambda_*)\frac{\operatorname{adj}(M(\sigma))}{\det(M(\sigma))} + O(\sigma-\lambda_*) = I - \frac{M'(\lambda_*)vu^H}{u^HM'(\lambda_*)v} + O(\sigma-\lambda_*), \quad \text{(A.3)}$$

where in the last step we expanded $\det(M(\sigma)) = \det(M(\lambda_*)) + (\sigma-\lambda_*)\frac{d}{d\lambda}\det(M(\lambda)) + O(\sigma-\lambda_*)^2$ and applied (A.2). This proves the formula for A_0 in the expansion (4.1).

We will derive the formula for A_1 by first noting that

$$A_1 = \left(\frac{d}{d\sigma}M(\lambda_*)M(\sigma)^{-1}\right)_{\sigma\to\lambda_*}.$$

Hence,

$$\frac{d}{d\sigma}M(\lambda_*)M(\sigma)^{-1} = -M(\lambda_*)M(\sigma)^{-1}M'(\sigma)M(\sigma)^{-1} =$$
$$-(A_0+(\sigma-\lambda_*)A_1)(M'(\lambda_*)+(\sigma-\lambda_*)M''(\lambda_*))\frac{(\operatorname{adj}(M(\lambda_*))+(\sigma-\lambda_*)\frac{d}{d\lambda_*}\operatorname{adj}(M(\lambda_*)))}{(\sigma-\lambda_*)\beta u^HM'(\lambda_*)v} + O(\sigma-\lambda_*),$$

where we inserted the (still unknown) expansion $M(\lambda_*)M(\sigma)^{-1} = A_0 + (\sigma-\lambda_*)A_1 + O(\sigma-\lambda_*)^2$. Note that $A_0M'(\lambda_*)\operatorname{adj}(M(\lambda_*)) = 0$. We now again use that $\operatorname{adj}(M(\lambda_*)) = \beta vu^H$ to find that

$$A_1 = \frac{-A_1M'(\lambda_*)\beta vu^H - A_0M''(\lambda_*)\beta vu^H - A_0M'(\lambda_*)\frac{d}{d\lambda_*}\operatorname{adj}(M(\lambda_*))}{\beta u^HM'(\lambda_*)v}$$

By solving for A_1, we have that

$$A_1 = \left(-\frac{A_0M'(\lambda_*)\frac{d}{d\lambda_*}\operatorname{adj}(M(\lambda_*))}{\beta u^HM'(\lambda_*)v} - \frac{A_0M''(\lambda_*)vu^H}{u^HM'(\lambda_*)v}\right)\left(I + \frac{M'(\lambda_*)vu^H}{u^HM'(\lambda_*)v}\right)^{-1}. \quad \text{(A.4)}$$

In the last step we will now compute the derivative of $M(\lambda_*)\operatorname{adj}(M(\sigma)) = M(\lambda_*)M(\sigma)^{-1}\det(M(\sigma))$. We find that,

$$M(\lambda_*)\left(\frac{d}{d\sigma}\operatorname{adj}(M(\sigma))\right)_{\sigma\to\lambda_*} = A_0(I\beta u^HM'(\lambda_*)v - \beta M'(\lambda_*)vu^H) = \beta(u^HM'(\lambda_*)v)A_0^2 = \beta(u^HM'(\lambda_*)v)A_0,$$

to which the potential solutions are parameterized by the variable $x\in\mathbb{C}^n$, such that

$$\left(\frac{d}{d\sigma}\operatorname{adj}(M(\sigma))\right)_{\sigma\to\lambda_*} = \beta(u^HM'(\lambda_*)v)M(\lambda_*)^+A_0 + vx^H. \quad \text{(A.5)}$$

Now note that the derivative of the adjoint matrix only appears in combination with the product

$$A_0M'(\lambda_*)\frac{d}{d\lambda_*}\operatorname{adj}(M(\lambda_*))$$

in (A.4). In this combination, the free variable term vx^H in (A.5) vanish since $A_0M'(\lambda_*)v = 0$. This fact and insertion of (A.5) into (A.4) completes the proof.

Acknowledgements This article presents results of the Belgian Programme on Interuniversity Poles of Attraction, initiated by the Belgian State, Prime Minister's Office for Science, Technology and Culture, the Optimization in Engineering Centre OPTEC of the K.U.Leuven, and the project STRT1-09/33 of the K.U.Leuven Research Foundation.

References

1. Anselone, P., Rall, L.: The solution of characteristic value-vector problems by Newton's method. Numer. Math. **11**, 38–45 (1968)
2. Betcke, M.: Iterative projection methods for symmetric nonlinear eigenvalue problems with applications. Ph.D. thesis, Technical University Hamburg-Harburg (2007)
3. Betcke, M., Voss, H.: Stationary Schrödinger equations governing electronic states of quantum dots in the presence of spinorbit splitting. Appl. Math. **52**, 267 – 284 (2007)
4. Betcke, T., Higham, N.J., Mehrmann, V., Schröder, C., Tisseur, F.: NLEVP: A collection of nonlinear eigenvalue problems. Tech. rep., Manchester Institute for Mathematical Sciences (2008)
5. Gohberg, I., Lancaster, P., Rodman, L.: Matrix polynomials. Academic press (1982)
6. Hadeler, K.: Mehrparametrige und nichtlineare Eigenwertaufgaben. Arch. Ration. Mech. Anal. **27**, 306–328 (1967)
7. Hryniv, R., Lancaster, P.: On the perturbation of analytic matrix functions. Integral Equations Oper. Theory **34**(3), 325–338 (1999)
8. Jarlebring, E.: The spectrum of delay-differential equations: numerical methods, stability and perturbation. Ph.D. thesis, TU Braunschweig (2008)
9. Jarlebring, E.: Convergence factors of Newton methods for nonlinear eigenvalue problems. Linear Algebra Appl. (2010). To appear
10. Jarlebring, E., Michiels, W.: Invariance properties in the root sensitivity of time-delay systems with double imaginary roots. Automatica **46**, 1112–1115 (2010)
11. Kressner, D.: A block Newton method for nonlinear eigenvalue problems. Numer. Math. **114**(2), 355–372 (2009)
12. Kublanovskaya, V.: On an approach to the solution of the generalized latent value problem for λ-matrices. SIAM J. Numer. Anal. **7**, 532–537 (1970)
13. Liao, B.S., Bai, Z., Lee, L.Q., Ko, K.: Nonlinear Rayleigh-Ritz iterative method for solving large scale nonlinear eigenvalue problems. Taiwanese Journal of Mathematics **14**(3), 869–883 (2010)
14. Meerbergen, K.: The quadratic Arnoldi method for the solution of the quadratic eigenvalue problem. SIAM J. Matrix Anal. Appl. **30**(4), 1463–1482 (2008)
15. Mehrmann, V., Voss, H.: Nonlinear eigenvalue problems: A challenge for modern eigenvalue methods. GAMM Mitteilungen **27**, 121–152 (2004)
16. Neumaier, A.: Residual inverse iteration for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. **22**, 914–923 (1985)
17. Peters, G., Wilkinson, J.: Inverse iterations, ill-conditioned equations and Newton's method. SIAM Rev. **21**, 339–360 (1979)
18. Rogers, E.: A minimax theory for overdamped systems. Arch. Ration. Mech. Anal. **16**, 89–96 (1964)
19. Rott, O., Jarlebring, E.: An iterative method for the multipliers of periodic delay-differential equations and the analysis of a PDE milling model. In: Proceedings of the 9th IFAC workshop on time-delay systems, Prague, pp. 1–6 (2010)
20. Ruhe, A.: Algorithms for the nonlinear eigenvalue problem. SIAM J. Numer. Anal. **10**, 674–689 (1973)
21. Schreiber, K.: Nonlinear eigenvalue problems: Newton-type methods and nonlinear Rayleigh functionals. Ph.D. thesis, TU Berlin (2008)
22. Schwetlick, H., Schreiber, K.: A Jacobi-Davidson like method for nonlinear eigenvalue problems based on singularity theory. PAMM **7**(1), 1020,401–1020,402 (2007)
23. Unger, H.: Nichtlineare Behandlung von Eigenwertaufgaben. Z. Angew. Math. Mech. **30**, 281–282 (1950). English translation: http://www.math.tu-dresden.de/~schwetli/Unger.html
24. Voss, H.: An Arnoldi method for nonlinear eigenvalue problems. BIT **44**, 387 – 401 (2004)
25. Voss, H.: Numerical methods for sparse nonlinear eigenvalue problems. In: Proc. XVth Summer School on Software and Algorithms of Numerical Mathematics, Hejnice, Czech Republic (2004). Report 70. Arbeitsbereich Mathematik, TU Hamburg-Harburg
26. Voss, H.: Iterative projection methods for computing relevant energy states of a quantum dot. J. Comput. Phys. **217**(2), 824–833 (2006)